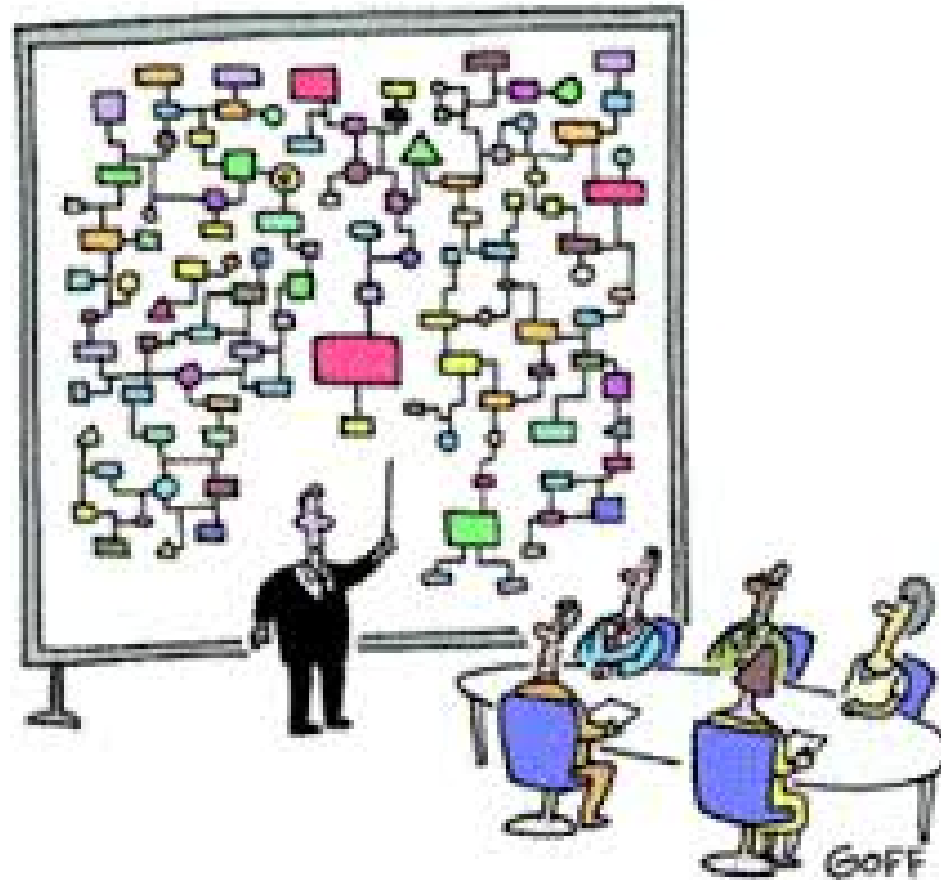


Bioinformatics

Week 14, 2020-06-02

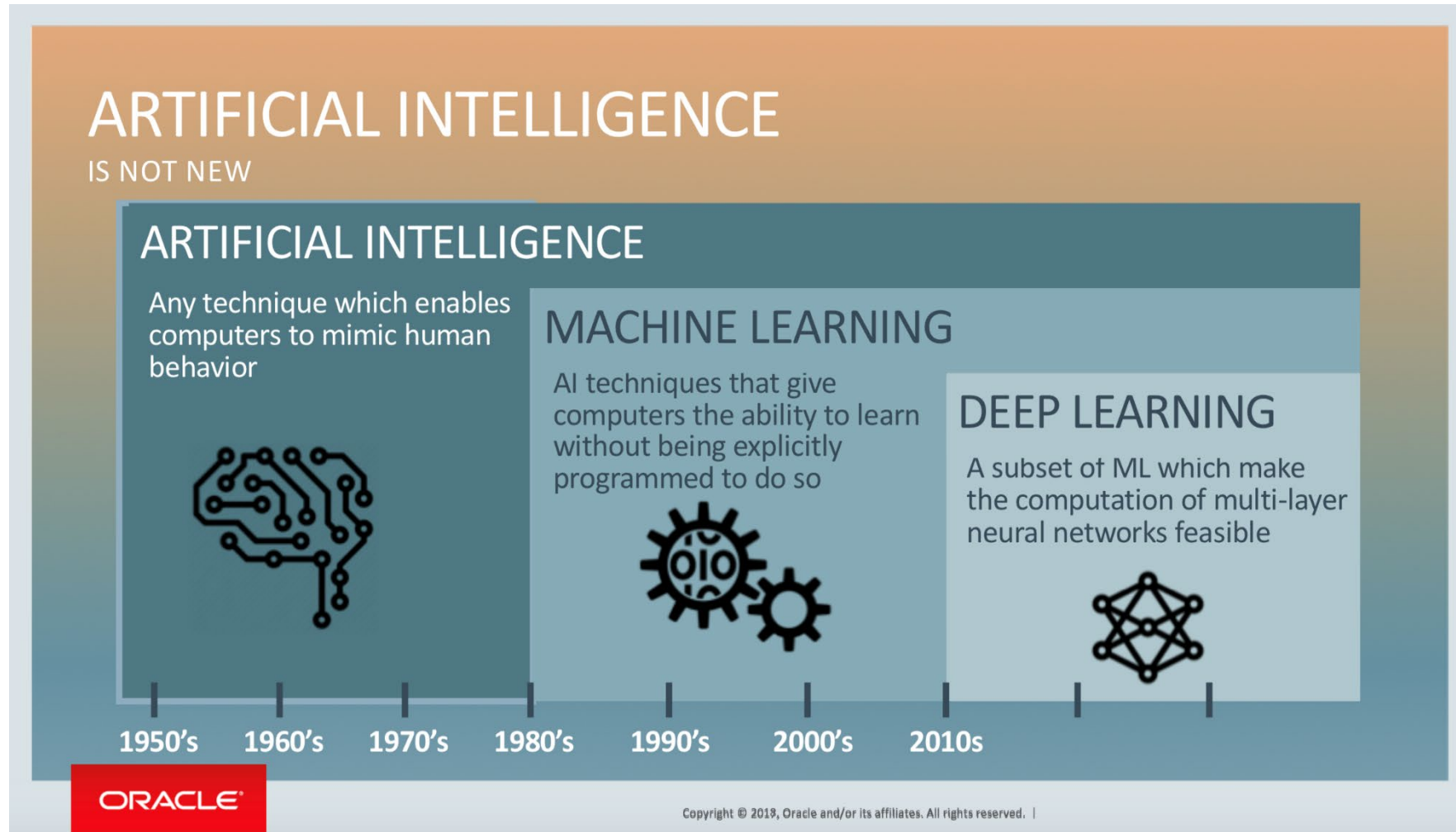
Xiaofan Zhou

Machine learning in biology



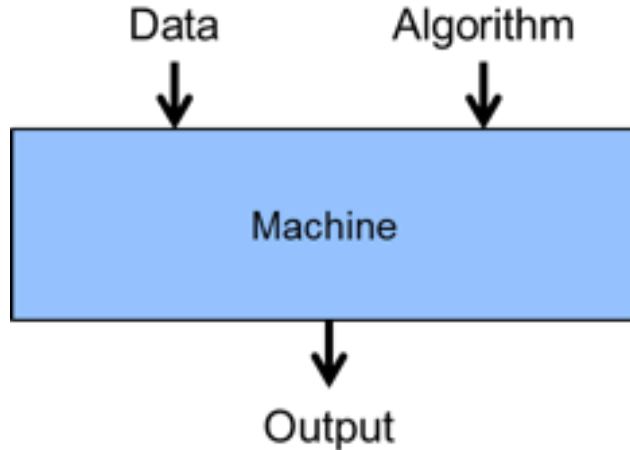
"And that's why we need a computer."

What is machine learning?

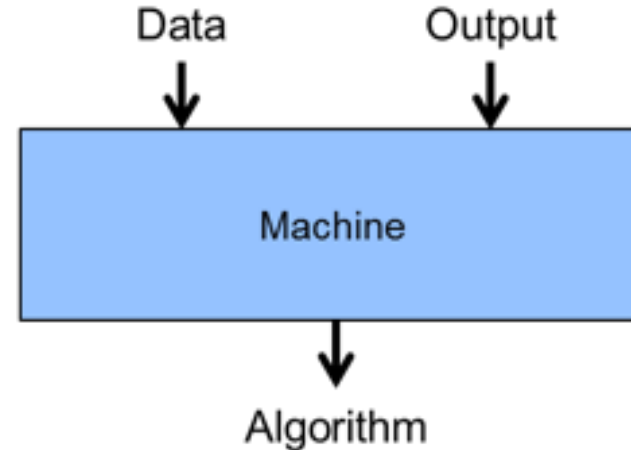


Traditional programming vs. machine learning

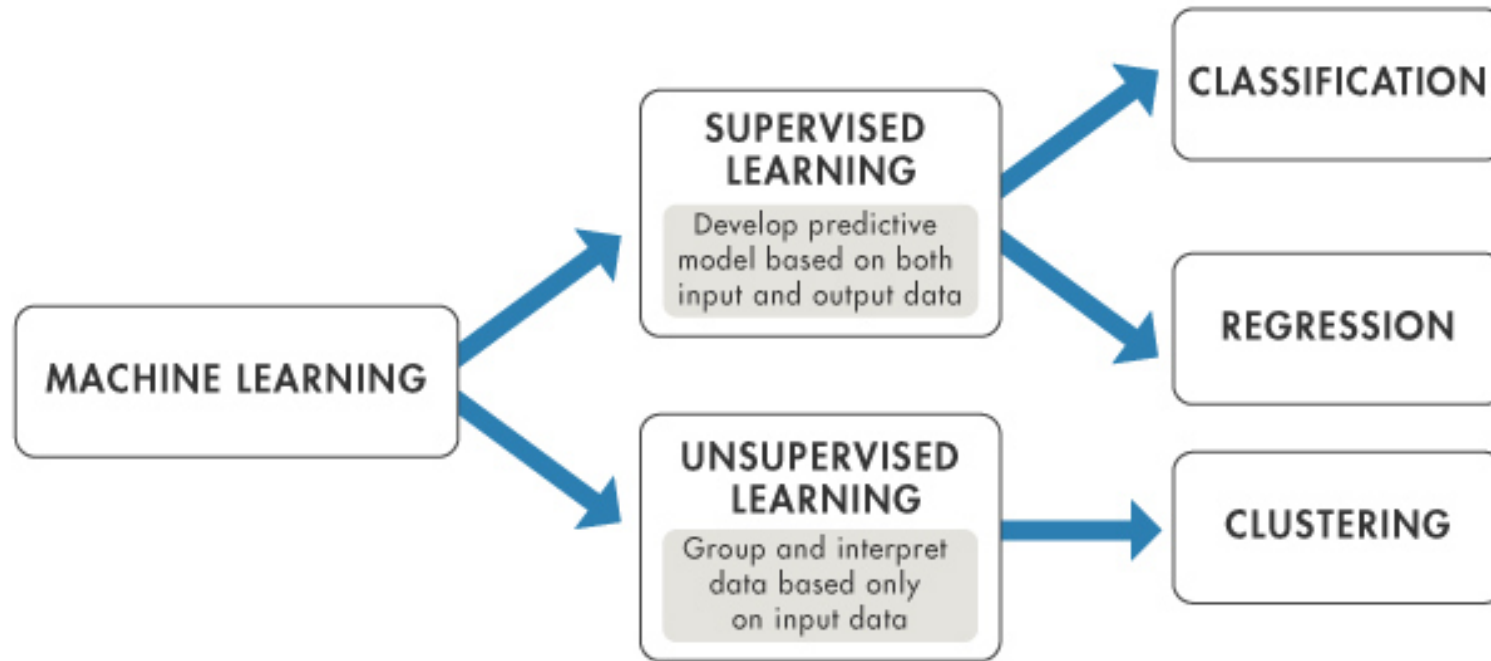
Traditional
programming



Machine
learning



Supervised vs. unsupervised machine learning



Machine learning vs. deep learning

Machine Learning



Deep Learning



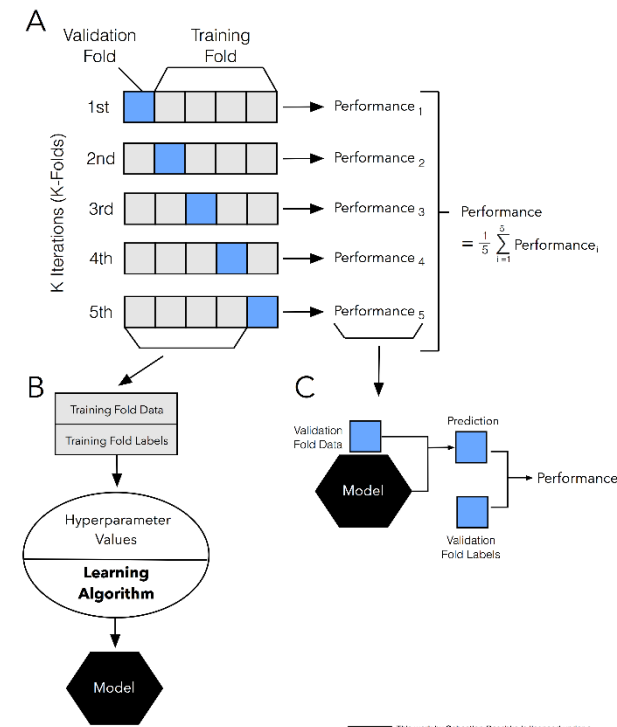
How does the machine learn?

From training to testing

1

Loading and
pre-processing
dataset of interest

Cross-validation



Deep Blue and AlphaGo



And AlphaStar on StarCraft II...



nature > articles > article a natureresearch journal

MENU nature

We'd like to understand how you use our websites in order to improve them. [Register your interest.](#)

Article | Published: 30 October 2019

Grandmaster level in StarCraft II using multi-agent reinforcement learning

Oriol Vinyals , Igor Babuschkin, [...] David Silver 

Nature 575, 350–354(2019) | [Cite this article](#)

54k Accesses | 21 Citations | 963 Altmetric | [Metrics](#)

Abstract

Many real-world applications require artificial agents to compete and coordinate with other agents in complex environments. As a stepping stone to this goal, the domain of StarCraft has emerged as an important challenge for artificial intelligence research, owing to its iconic and enduring status among the most difficult professional esports and its relevance to the real world in terms of its raw complexity and multi-agent challenges. Over the course of a decade and numerous competitions^{1,2,3}, the strongest agents have simplified important aspects of the game, utilized superhuman capabilities, or employed hand-crafted sub-

Associated Content

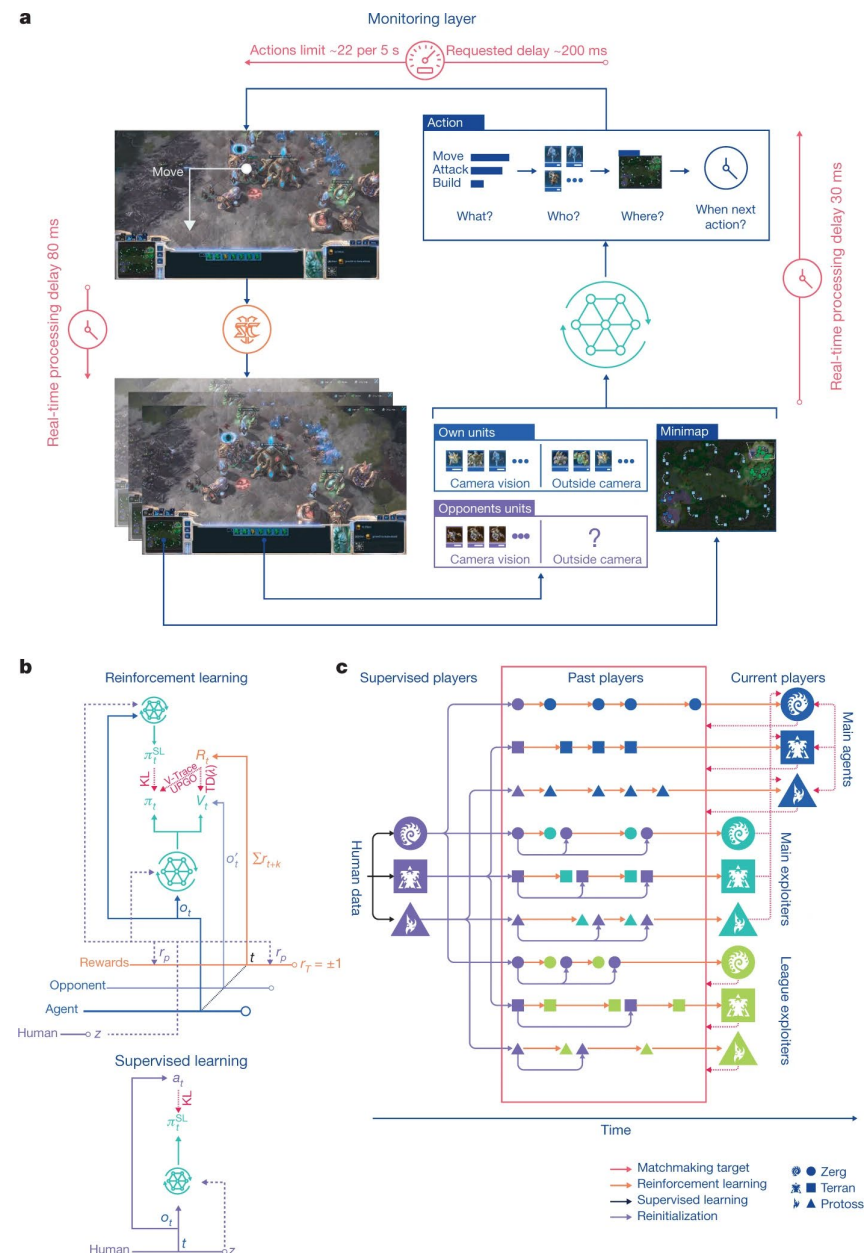
Nature | Nature Video

DeepMind's new AI masters the online game StarCraft II

Thom Hoffman

Sections Figures References

Abstract
Main
Learning algorithm
Empirical evaluation
Conclusion
Methods
Data availability
Code availability
References
Acknowledgements
Author information



AlphaFold for protein folding

MENU ▾

nature

Search

E-alert


Submit

Login


We'd like to understand how you use our websites in order to improve them. [Register your interest](#)

Article | Published: 15 January 2020

Improved protein structure prediction using potentials from deep learning

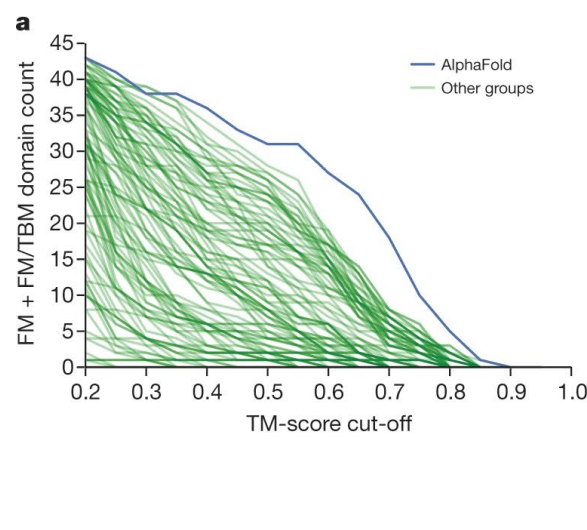
Andrew W. Senior , Richard Evans, John Jumper, James Kirkpatrick, Laurent Sifre, Tim Green, Chongli Qin, Augustin Židek, Alexander W. R. Nelson, Alex Bridgland, Hugo Penedones, Stig Petersen, Karen Simonyan, Steve Crossan, Pushmeet Kohli, David T. Jones, David Silver, Koray Kavukcuoglu & Demis Hassabis

You have full access to this article via
South China Agricultural University Library

[Download PDF](#) 

Associated Content

a



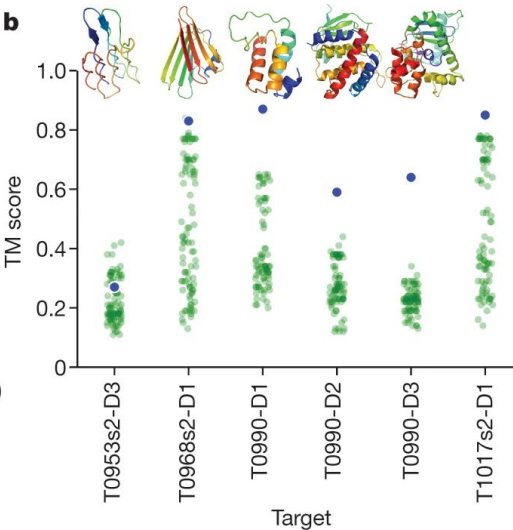
FM + FM/TBM domain count

TM-score cut-off

— AlphaFold

— Other groups

b

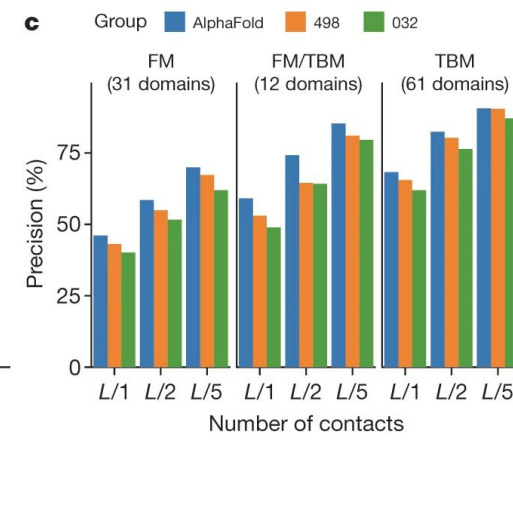


TM score

Target

T0953s2-D3 T0968s2-D1 T0990-D1 T0990-D2 T0990-D3 T1017s2-D1

c



Precision (%)

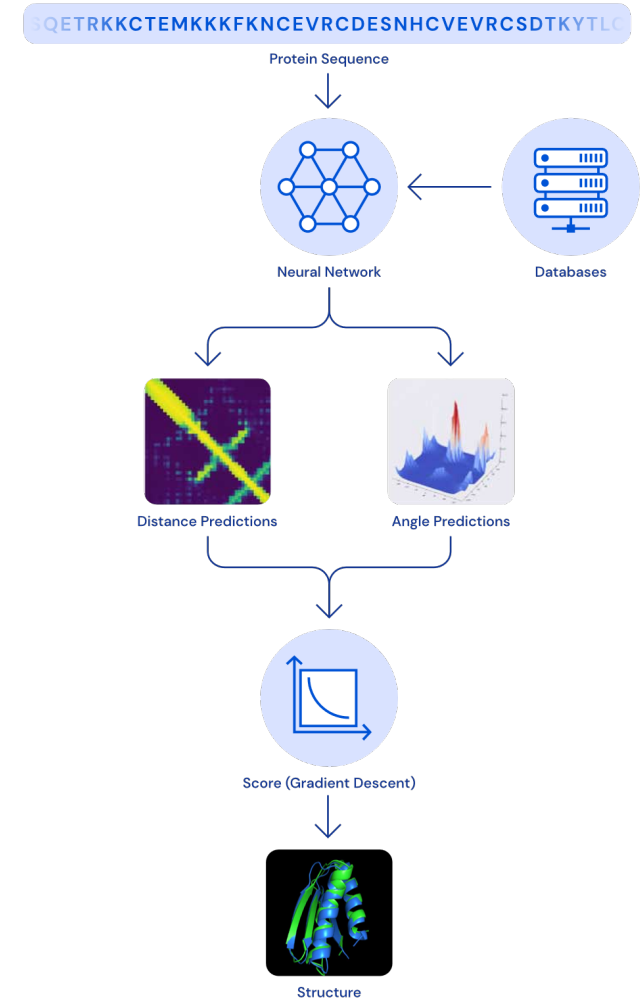
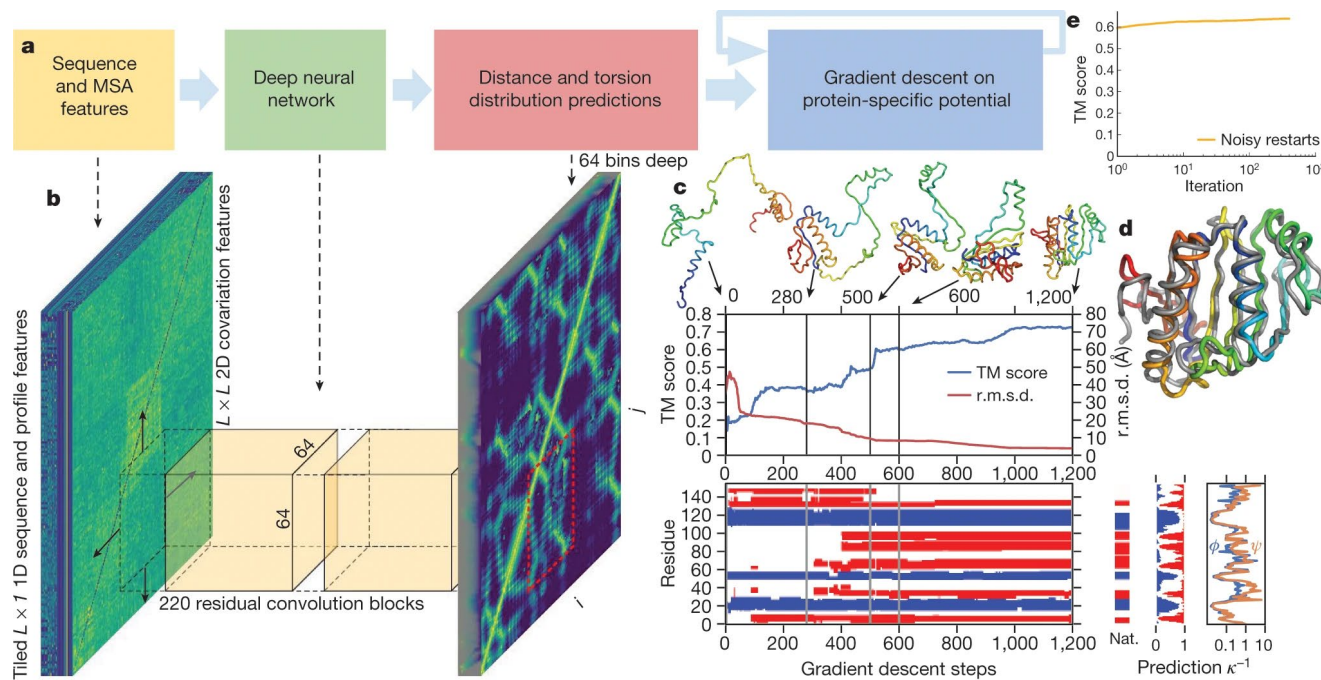
Number of contacts

Group AlphaFold 498 032

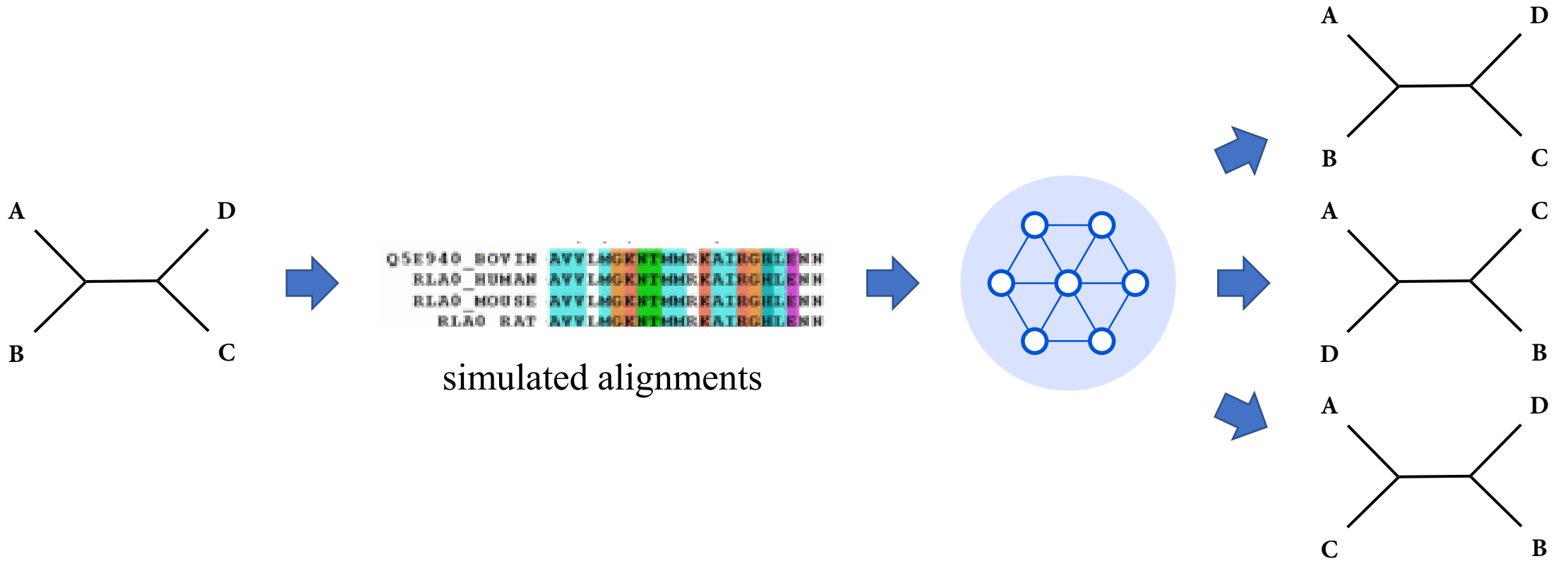
FM (31 domains) FM/TBM (12 domains) TBM (61 domains)

L/1 L/2 L/5 L/1 L/2 L/5 L/1 L/2 L/5

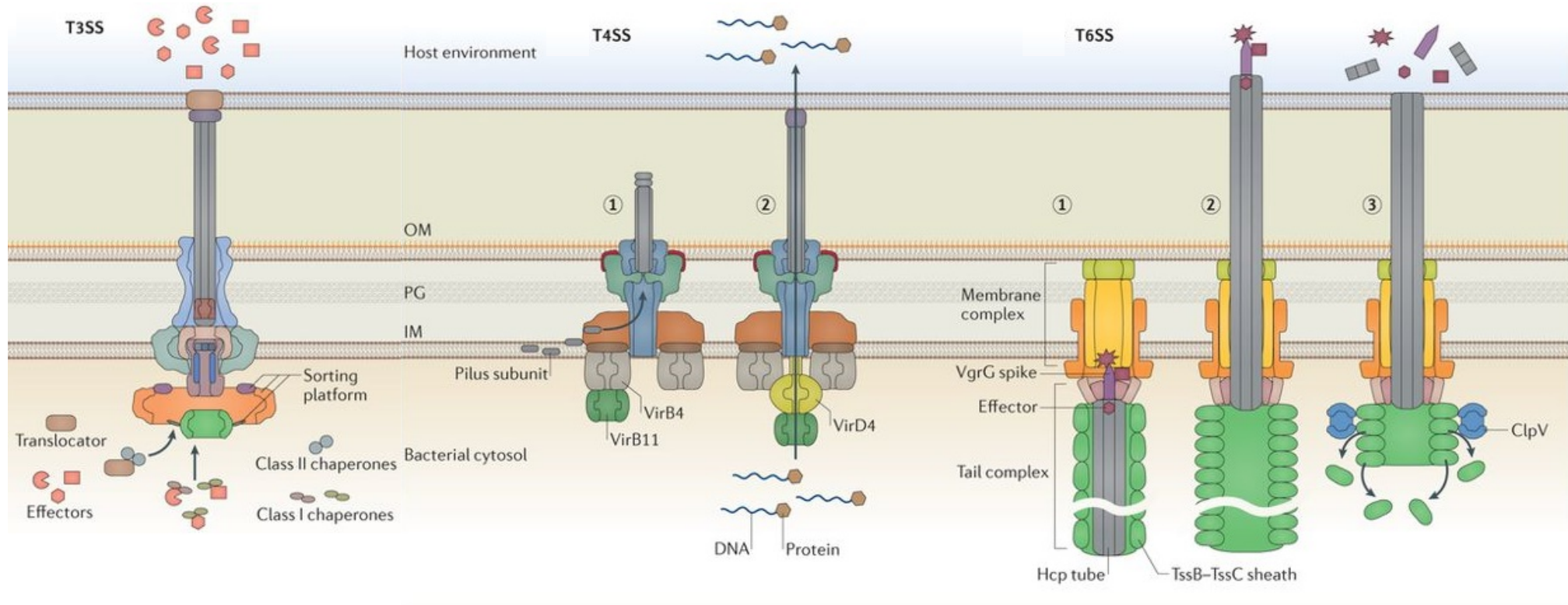
AlphaFold for protein folding



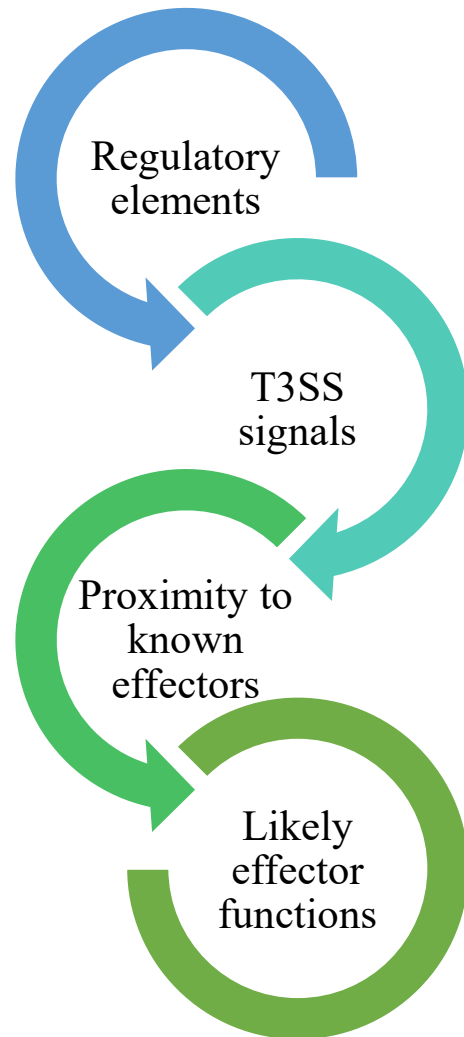
CNN/PhyDL for phylogenetics



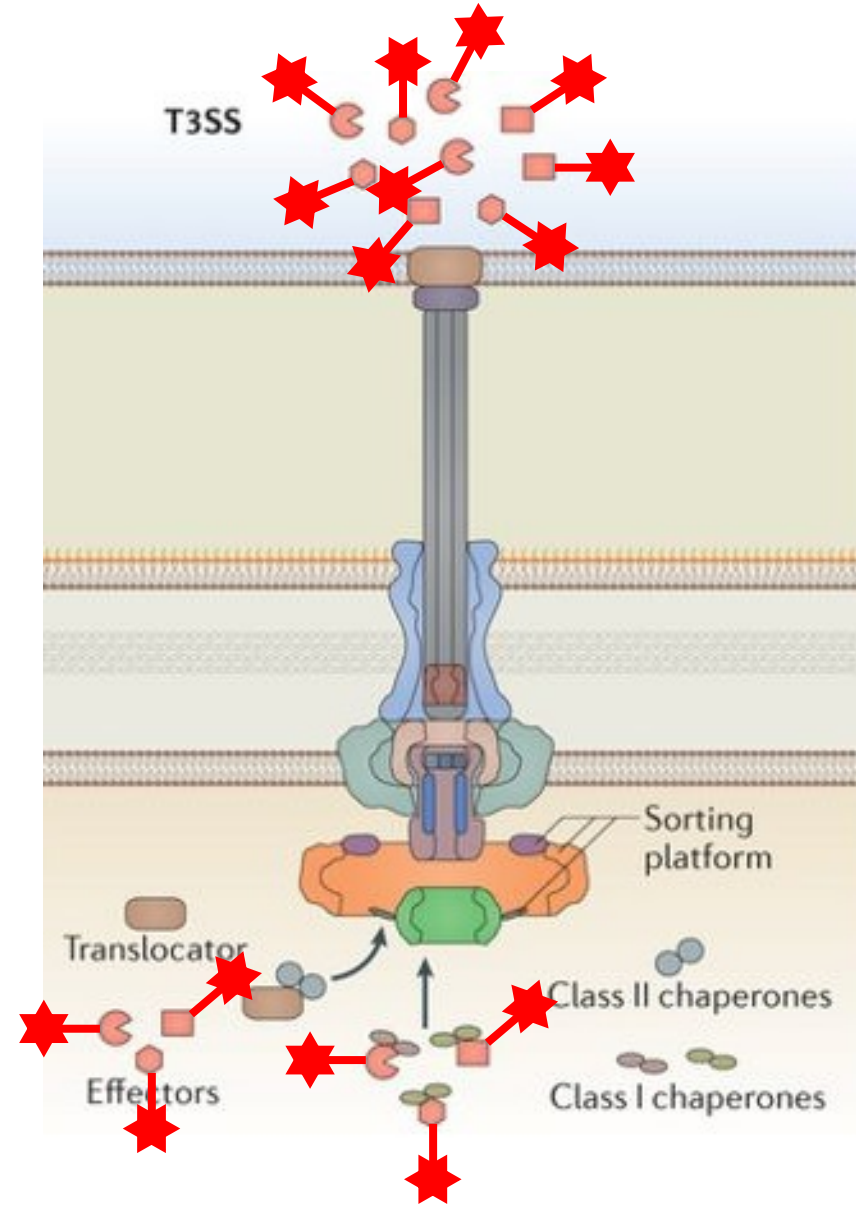
Bacterial secretion systems and effectors



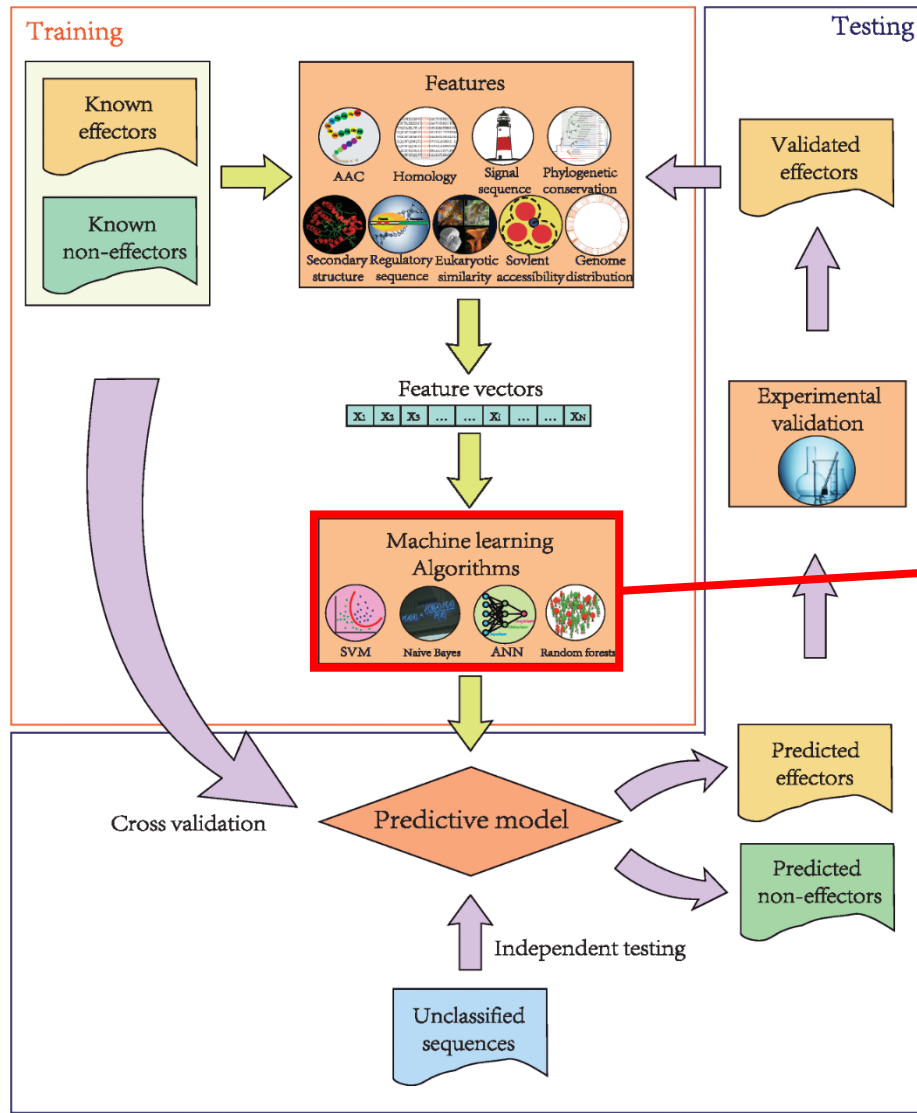
Identification of effectors: *experimental approach*



★
in vivo reporter
(e.g., adenylate
cyclase, AvrRpt2)

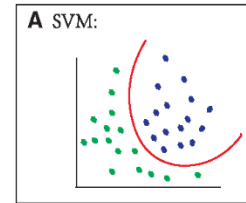
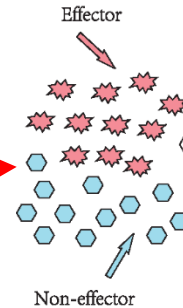


Identification of effectors: *bioinformatics approach*



Input data

Prediction approaches

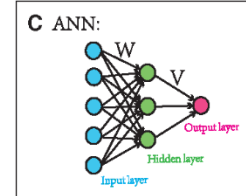


Supporting Vector Machine

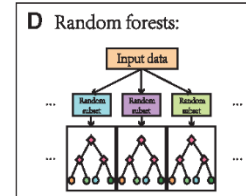
B Naive Bayes:

$$p(c|x) = \frac{p(x|c) p(c)}{p(x)}$$
$$= \frac{\prod_i p(x_i|c) p(c)}{\prod_i p(x_i)}$$

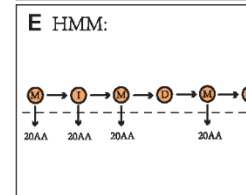
Naïve Bayes



Artificial Neural Network

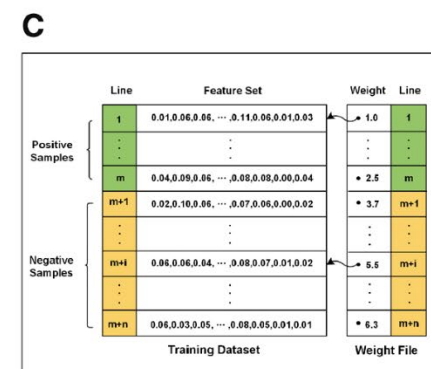
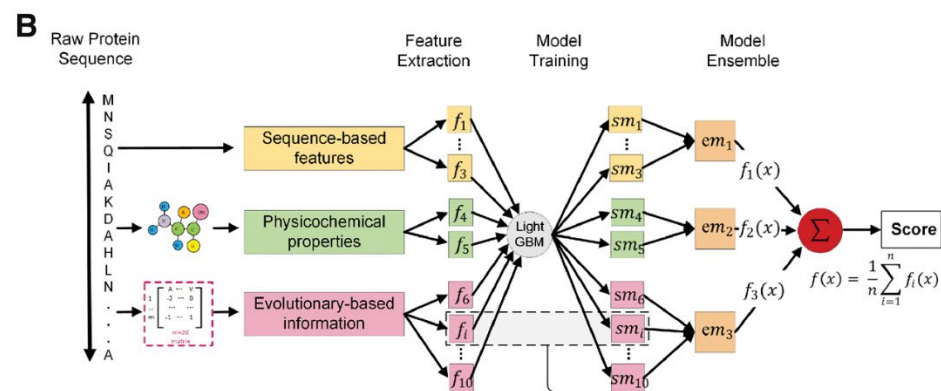
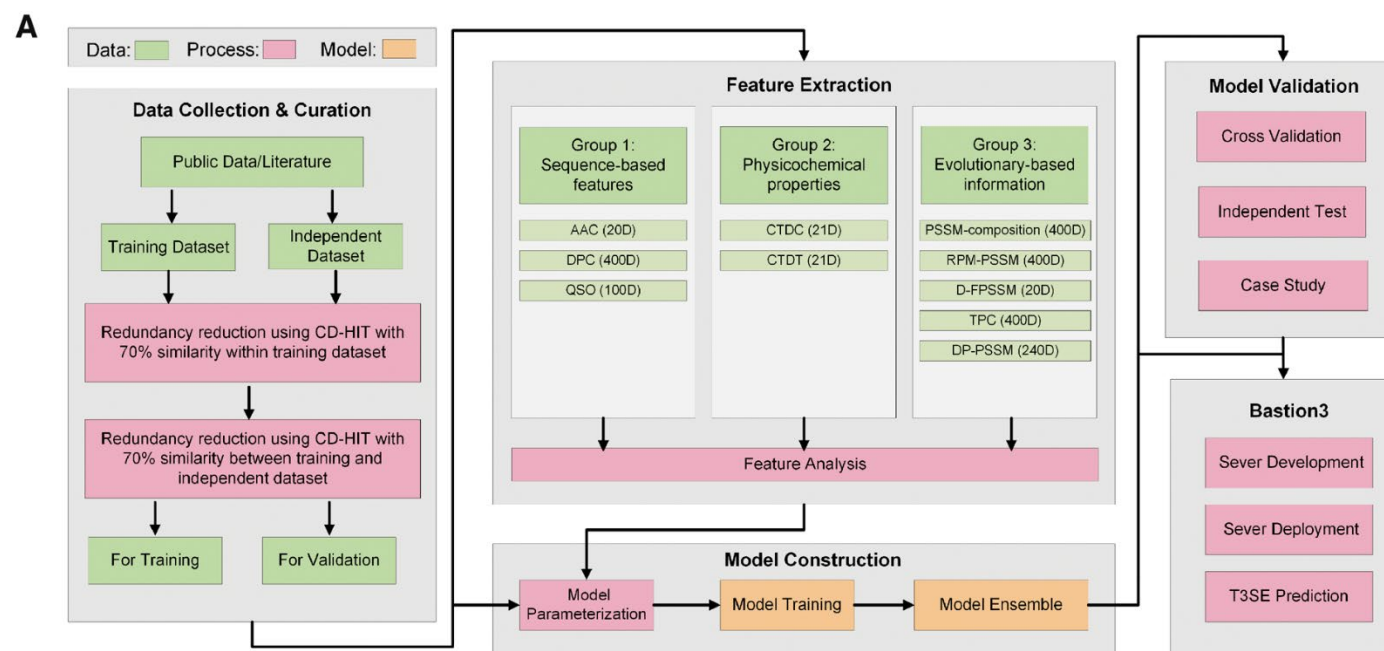


Random Forests

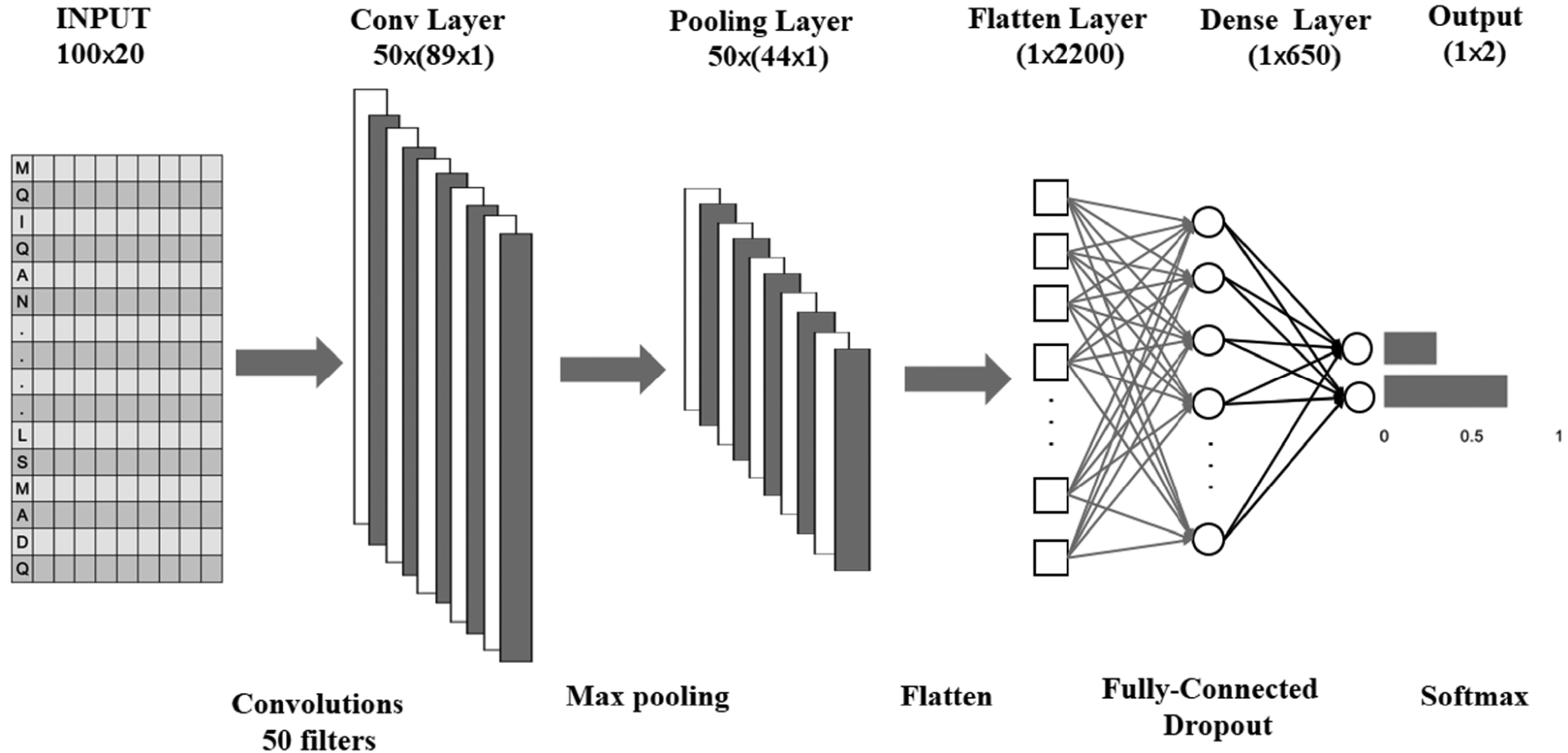


Hidden Markov Model

Design of a prediction method: Bastion 3

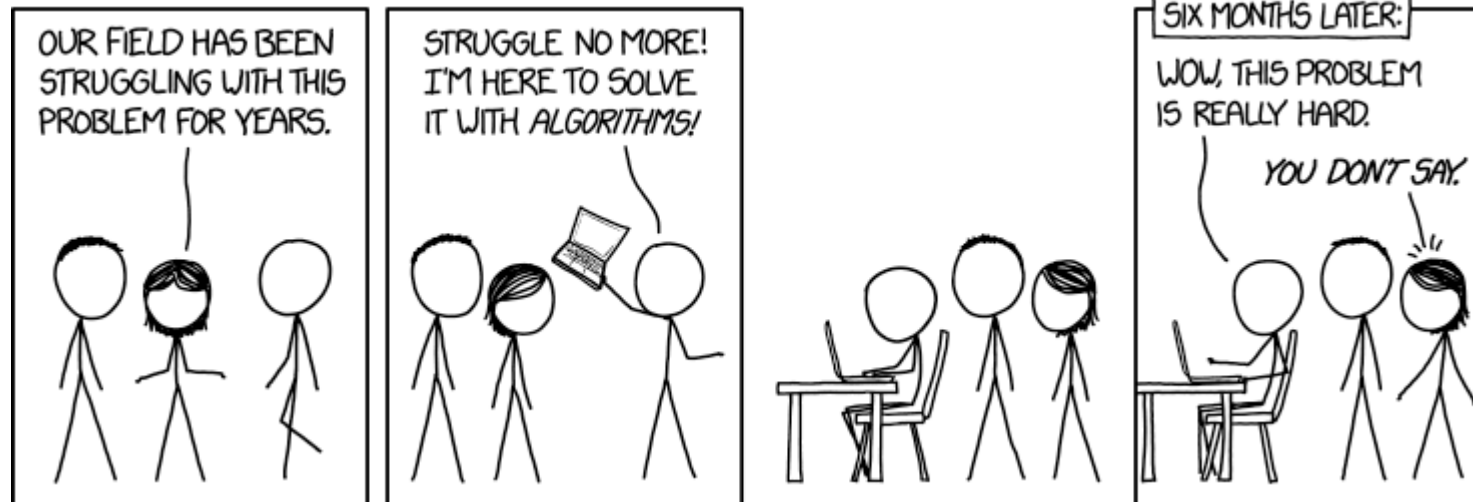


Design of a prediction method: DeepT3



Many other applications

- Genetic variation
- Gene annotation
- Gene function prediction
- Protein subcellular localization
- ...



Questions?